Locally Equivalent Weights for Bayesian MrP

Ryan Giordano, Alice Cima, Erin Hartman, Jared Murray, Avi Feller
University of British Columbia Statistics Seminar
October 2025

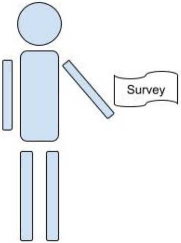
    

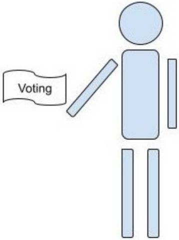
The basic problem

We have a survey population, for whom we observe:

- Covariates $\mathbf{x}$ (e.g. race, gender, zip code, age, education level)
- Responses y (e.g. A binary response to "do you support Trump")

We want the average response in a target population, in which we observe only covariates.





Observe $(\mathbf{x}_i, y_i)$ for $i = 1, \dots, N_S$ | Observe $\mathbf{x}_j$ for $j = 1, \dots, N_T$

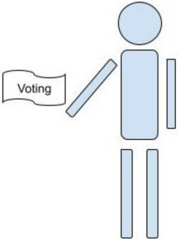
[1]Photo copyright: Mark Taylor / `naturepl.com`

The basic problem

We have a survey population, for whom we observe:

- Covariates $\mathbf{x}$ (e.g. race, gender, zip code, age, education level)
- Responses y (e.g. A binary response to "do you support Trump")

We want the average response in a target population, in which we observe only covariates.



Observe $(\mathbf{x}_i, y_i)$ for $i = 1, \ldots, N_S$ | Observe $\mathbf{x}_j$ for $j = 1, \ldots, N_T$

The problem is that the populations may be very different, maybe leading to bias. [1]

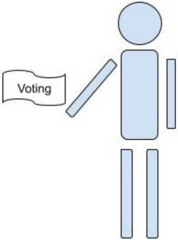
[1] Photo copyright: Mark Taylor / `naturepl.com`

The basic problem

We have a survey population, for whom we observe:

- Covariates **x** (e.g. race, gender, zip code, age, education level)
- Responses y (e.g. A binary response to "do you support Trump")

We want the average response in a target population, in which we observe only covariates.



Observe $(\mathbf{x}_i, y_i)$ for $i = 1, \ldots, N_S$ | Observe $\mathbf{x}_j$ for $j = 1, \ldots, N_T$

The problem is that the populations may be very different, maybe leading to bias. [1]

How can we use the covariates to say something about the target responses?

[1] Photo copyright: Mark Taylor / `naturepl.com`

Two approaches

We want $\mu := \frac{1}{N_T} \sum_{j=1}^{N_T} y_j$, but don't observe target y_j. Let $Y_{\mathcal{S}} = \{y_1, \ldots, y_{N_S}\}$.

- Assume $p(y|\mathbf{x})$ is the same in both populations,
- But the distribution of **x** may be different in the survey and target.

Two approaches

We want $\mu := \frac{1}{N_T} \sum_{j=1}^{N_T} y_j$, but don't observe target y_j. Let $Y_{\mathcal{S}} = \{y_1, \ldots, y_{N_S}\}$.

- Assume $p(y|\mathbf{x})$ is the same in both populations,
- But the distribution of $\mathbf{x}$ may be different in the survey and target.

Calibration weighting (CW)

▶ Choose "calibration weights" w_i
using only the regressors $\mathbf{x}$
(e.g. raking weights)

Bayesian hierarchical modeling (MrP)

▶ Choose $\mathbb{E}\left[y|\mathbf{x}, \theta\right] = m(\theta^\intercal \mathbf{x})$,
choose prior $\mathcal{P}(\theta|\Sigma)\mathcal{P}(\Sigma)$
(e.g. Hierarchical logistic regression)

Two approaches

We want $\mu := \frac{1}{N_T} \sum_{j=1}^{N_T} y_j$, but don't observe target y_j. Let $Y_{\mathcal{S}} = \{y_1, \ldots, y_{N_S}\}$.

- Assume $p(y|\mathbf{x})$ is the same in both populations,
- But the distribution of $\mathbf{x}$ may be different in the survey and target.

Calibration weighting (CW)

▶ Choose "calibration weights" w_i using only the regressors $\mathbf{x}$ (e.g. raking weights)

▶ Take $\hat{\mu}^{\text{CW}}(Y_{\mathcal{S}}) = \frac{1}{N_S} \sum_{i=1}^{N_S} w_i y_i$

Bayesian hierarchical modeling (MrP)

▶ Choose $\mathbb{E}\left[y|\mathbf{x}, \theta\right] = m(\theta^\intercal \mathbf{x})$, choose prior $\mathcal{P}(\theta|\Sigma)\mathcal{P}(\Sigma)$ (e.g. Hierarchical logistic regression)

▶ Take $\hat{y}_j = \mathbb{E}_{\mathcal{P}(\theta|\text{Survey data})}\left[y|\mathbf{x}_j\right]$ and $\hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T} \sum_{j=1}^{N_T} \hat{y}_j$

Two approaches

We want $\mu := \frac{1}{N_T} \sum_{j=1}^{N_T} y_j$, but don't observe target y_j. Let $Y_{\mathcal{S}} = \{y_1, \ldots, y_{N_S}\}$.

- Assume $p(y|\mathbf{x})$ is the same in both populations,
- But the distribution of $\mathbf{x}$ may be different in the survey and target.

Calibration weighting (CW)

- ▶ Choose "calibration weights" w_i using only the regressors $\mathbf{x}$ (e.g. raking weights)
- ▶ Take $\hat{\mu}^{\text{CW}}(Y_{\mathcal{S}}) = \frac{1}{N_S} \sum_{i=1}^{N_S} w_i y_i$
- ▶ Dependence on y_i is clear

Bayesian hierarchical modeling (MrP)

- ▶ Choose $\mathbb{E}\left[y|\mathbf{x}, \theta\right] = m(\theta^\intercal \mathbf{x})$, choose prior $\mathcal{P}(\theta|\Sigma)\mathcal{P}(\Sigma)$ (e.g. Hierarchical logistic regression)
- ▶ Take $\hat{y}_j = \mathbb{E}_{\mathcal{P}(\theta|\text{Survey data})}\left[y|\mathbf{x}_j\right]$ and $\hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T} \sum_{j=1}^{N_T} \hat{y}_j$
- ▶ Dependence on y_i very complicated (Typically via MCMC draws from $\mathcal{P}(\theta|\text{Survey data})$)

Two approaches

We want $\mu := \frac{1}{N_T} \sum_{j=1}^{N_T} y_j$, but don't observe target y_j. Let $Y_{\mathcal{S}} = \{y_1, \ldots, y_{N_S}\}$.

- Assume $p(y|\mathbf{x})$ is the same in both populations,
- But the distribution of $\mathbf{x}$ may be different in the survey and target.

Calibration weighting (CW)

- ▶ Choose "calibration weights" w_i using only the regressors $\mathbf{x}$ (e.g. raking weights)
- ▶ Take $\hat{\mu}^{\text{CW}}(Y_{\mathcal{S}}) = \frac{1}{N_S} \sum_{i=1}^{N_S} w_i y_i$
- ▶ Dependence on y_i is clear
- ▶ Weights give interpretable diagnostics:
  - Regressor balance
  - Frequentist variability
  - Partial pooling
  - Extraplolation

Bayesian hierarchical modeling (MrP)

- ▶ Choose $\mathbb{E}\left[y|\mathbf{x}, \theta\right] = m(\theta^\intercal \mathbf{x})$, choose prior $\mathcal{P}(\theta|\Sigma)\mathcal{P}(\Sigma)$ (e.g. Hierarchical logistic regression)
- ▶ Take $\hat{y}_j = \mathbb{E}_{\mathcal{P}(\theta|\text{Survey data})}\left[y|\mathbf{x}_j\right]$ and $\hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T} \sum_{j=1}^{N_T} \hat{y}_j$
- ▶ Dependence on y_i very complicated (Typically via MCMC draws from $\mathcal{P}(\theta|\text{Survey data})$)
- ▶ **Black box**

Two approaches

We want $\mu := \frac{1}{N_T} \sum_{j=1}^{N_T} y_j$, but don't observe target y_j. Let $Y_{\mathcal{S}} = \{y_1, \ldots, y_{N_S}\}$.

- Assume $p(y|\mathbf{x})$ is the same in both populations,
- But the distribution of $\mathbf{x}$ may be different in the survey and target.

Calibration weighting (CW)

- ▶ Choose "calibration weights" w_i using only the regressors $\mathbf{x}$ (e.g. raking weights)
- ▶ Take $\hat{\mu}^{\mathrm{CW}}(Y_{\mathcal{S}}) = \frac{1}{N_S} \sum_{i=1}^{N_S} w_i y_i$
- ▶ Dependence on y_i is clear
- ▶ Weights give interpretable diagnostics:
  - Regressor balance
  - Frequentist variability
  - Partial pooling
  - Extraplolation

Bayesian hierarchical modeling (MrP)

- ▶ Choose $\mathbb{E}\left[y|\mathbf{x}, \theta\right] = m(\theta^\intercal \mathbf{x})$, choose prior $\mathcal{P}(\theta|\Sigma)\mathcal{P}(\Sigma)$ (e.g. Hierarchical logistic regression)
- ▶ Take $\hat{y}_j = \mathbb{E}_{\mathcal{P}(\theta|\text{Survey data})}\left[y|\mathbf{x}_j\right]$ and $\hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T} \sum_{j=1}^{N_T} \hat{y}_j$
- ▶ Dependence on y_i very complicated (Typically via MCMC draws from $\mathcal{P}(\theta|\text{Survey data})$)
- ▶ **Black box**

$\leftarrow$ We open this box, providing analogues of all these diagnostics

Prior work: Equivalent weights for linear models

Gelman (2007b) observes that MrP is a CW estimator when one uses linear regression to form $\hat{y}$:

$$\hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T}\sum_{j=1}^{N_T}\hat{y}_j = \frac{1}{N_T}\sum_{j=1}^{N_T}\underbrace{\mathbf{x}_j^\intercal\hat{\theta}}_{\text{Linear in } Y_{\mathcal{S}}}$$

Most existing literature on comparing CW and MrP focus on such linear models. [2]

[2] For example, Gelman (2007b), B., F., and H. (2021), and Chattopadhyay and Zubizarreta (2023).

Prior work: Equivalent weights for linear models

Gelman (2007b) observes that MrP is a CW estimator when one uses linear regression to form $\hat{y}$:

$$\hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T}\sum_{j=1}^{N_T} \hat{y}_j = \frac{1}{N_T}\sum_{j=1}^{N_T} \underbrace{\mathbf{x}_j^\intercal \hat{\theta}}_{\text{Linear in } Y_{\mathcal{S}}}$$

Most existing literature on comparing CW and MrP focus on such linear models. [2]

But what if you use a non–linear link function? Or a hierarchical model?

> *"It would also be desirable to use nonlinear methods ... but then it would seem difficult to construct even approximately equivalent weights. Weighting and fully nonlinear models would seem to be completely incompatible methods." — (Gelman 2007a)*

[2] For example, Gelman (2007b), B., F., and H. (2021), and Chattopadhyay and Zubizarreta (2023).

Approximately equivalent weights for (some) logistic regression MrP

- Suppose the model is $m(\mathbf{x}^\intercal\theta) = \mathrm{Logistic}(\mathbf{x}^\intercal\theta)$, with MLE $\hat{\theta}$.
- MrP is $\hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T}\sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal\hat{\theta})$.

Approximately equivalent weights for (some) logistic regression MrP

- Suppose the model is $m(\mathbf{x}^\intercal \theta) = \text{Logistic}(\mathbf{x}^\intercal \theta)$, with MLE $\hat{\theta}$.
- MrP is $\hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T} \sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal \hat{\theta})$.

The map from $Y_{\mathcal{S}} \mapsto m(\mathbf{x}_j^\intercal \hat{\theta})$ is *inherently nonlinear*.

But *some sample averages* of $m(\mathbf{x}_j^\intercal \hat{\theta})$ can be approximately linear.

Approximately equivalent weights for (some) logistic regression MrP

- Suppose the model is $m(\mathbf{x}^\intercal \theta) = \mathrm{Logistic}(\mathbf{x}^\intercal \theta)$, with MLE $\hat{\theta}$.
- MrP is $\hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T} \sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal \hat{\theta})$.

The map from $Y_{\mathcal{S}} \mapsto m(\mathbf{x}_j^\intercal \hat{\theta})$ is *inherently nonlinear*.

But *some sample averages* of $m(\mathbf{x}_j^\intercal \hat{\theta})$ can be approximately linear.

Example

Suppose $\frac{\mathcal{P}_T(\mathbf{x})}{\mathcal{P}_S(\mathbf{x})} \approx \alpha^\intercal \mathbf{x}$ for some α. **Then MrP is a *approximately* a CW estimator.**

Approximately equivalent weights for (some) logistic regression MrP

- Suppose the model is $m(\mathbf{x}^\intercal\theta) = \mathrm{Logistic}(\mathbf{x}^\intercal\theta)$, with MLE $\hat{\theta}$.
- MrP is $\hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T}\sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal\hat{\theta})$.

Example

Suppose $\frac{\mathcal{P}_T(\mathbf{x})}{\mathcal{P}_S(\mathbf{x})} \approx \alpha^\intercal\mathbf{x}$ for some α. **Then MrP is a *approximately* a CW estimator.**

Approximately equivalent weights for (some) logistic regression MrP

- Suppose the model is $m(\mathbf{x}^\intercal \theta) = \mathrm{Logistic}(\mathbf{x}^\intercal \theta)$, with MLE $\hat{\theta}$.
- MrP is $\hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T} \sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal \hat{\theta})$.

Example

Suppose $\frac{\mathcal{P}_T(\mathbf{x})}{\mathcal{P}_S(\mathbf{x})} \approx \alpha^\intercal \mathbf{x}$ for some α. **Then MrP is a *approximately* a CW estimator.**

$$\hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T} \sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal \hat{\theta})$$

Approximately equivalent weights for (some) logistic regression MrP

- Suppose the model is $m(\mathbf{x}^\intercal \theta) = \mathrm{Logistic}(\mathbf{x}^\intercal \theta)$, with MLE $\hat{\theta}$.
- MrP is $\hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T} \sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal \hat{\theta})$.

Example

Suppose $\frac{\mathcal{P}_T(\mathbf{x})}{\mathcal{P}_S(\mathbf{x})} \approx \alpha^\intercal \mathbf{x}$ for some α. **Then MrP is a *approximately* a CW estimator.**

$$
\begin{aligned}
\hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}}) &= \frac{1}{N_T} \sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal \hat{\theta}) \\
&\approx \int m(\mathbf{x}^\intercal \hat{\theta}) \mathcal{P}_T(\mathbf{x}) d\mathbf{x} && \text{(Law of large numbers)}
\end{aligned}
$$

Approximately equivalent weights for (some) logistic regression MrP

- Suppose the model is $m(\mathbf{x}^\intercal\theta) = \mathrm{Logistic}(\mathbf{x}^\intercal\theta)$, with MLE $\hat{\theta}$.
- MrP is $\hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T}\sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal\hat{\theta})$.

Example

Suppose $\frac{\mathcal{P}_T(\mathbf{x})}{\mathcal{P}_S(\mathbf{x})} \approx \alpha^\intercal\mathbf{x}$ for some α. **Then MrP is a *approximately* a CW estimator.**

$$
\begin{aligned}
\hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}}) &= \frac{1}{N_T}\sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal\hat{\theta}) \\
&\approx \int m(\mathbf{x}^\intercal\hat{\theta})\mathcal{P}_T(\mathbf{x})d\mathbf{x} && \text{(Law of large numbers)} \\
&= \int \frac{\mathcal{P}_T(\mathbf{x})}{\mathcal{P}_S(\mathbf{x})} m(\mathbf{x}^\intercal\hat{\theta})\mathcal{P}_S(\mathbf{x})d\mathbf{x} && \text{(Multiply by } \mathcal{P}_S(\mathbf{x})/\mathcal{P}_S(\mathbf{x})\text{)}
\end{aligned}
$$

Approximately equivalent weights for (some) logistic regression MrP

- Suppose the model is $m(\mathbf{x}^\intercal \theta) = \mathrm{Logistic}(\mathbf{x}^\intercal \theta)$, with MLE $\hat{\theta}$.
- MrP is $\hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T} \sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal \hat{\theta})$.

Example

Suppose $\frac{\mathcal{P}_T(\mathbf{x})}{\mathcal{P}_S(\mathbf{x})} \approx \alpha^\intercal \mathbf{x}$ for some α. **Then MrP is a *approximately* a CW estimator.**

$$
\begin{aligned}
\hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}}) &= \frac{1}{N_T} \sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal \hat{\theta}) \\
&\approx \int m(\mathbf{x}^\intercal \hat{\theta}) \mathcal{P}_T(\mathbf{x}) d\mathbf{x} && \text{(Law of large numbers)} \\
&= \int \frac{\mathcal{P}_T(\mathbf{x})}{\mathcal{P}_S(\mathbf{x})} m(\mathbf{x}^\intercal \hat{\theta}) \mathcal{P}_S(\mathbf{x}) d\mathbf{x} && \text{(Multiply by } \mathcal{P}_S(\mathbf{x})/\mathcal{P}_S(\mathbf{x})) \\
&\approx \int (\alpha^\intercal \mathbf{x})\, m(\mathbf{x}^\intercal \hat{\theta}) \mathcal{P}_S(\mathbf{x}) d\mathbf{x} && \text{(By assumption)}
\end{aligned}
$$

Approximately equivalent weights for (some) logistic regression MrP

- Suppose the model is $m(\mathbf{x}^\intercal \theta) = \mathrm{Logistic}(\mathbf{x}^\intercal \theta)$, with MLE $\hat{\theta}$.
- MrP is $\hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T} \sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal \hat{\theta})$.

Example

Suppose $\frac{\mathcal{P}_T(\mathbf{x})}{\mathcal{P}_S(\mathbf{x})} \approx \alpha^\intercal \mathbf{x}$ for some α. **Then MrP is a *approximately* a CW estimator.**

$$
\begin{aligned}
\hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}}) &= \frac{1}{N_T} \sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal \hat{\theta}) \\
&\approx \int m(\mathbf{x}^\intercal \hat{\theta}) \mathcal{P}_T(\mathbf{x}) d\mathbf{x} && \text{(Law of large numbers)} \\
&= \int \frac{\mathcal{P}_T(\mathbf{x})}{\mathcal{P}_S(\mathbf{x})} m(\mathbf{x}^\intercal \hat{\theta}) \mathcal{P}_S(\mathbf{x}) d\mathbf{x} && \text{(Multiply by } \mathcal{P}_S(\mathbf{x})/\mathcal{P}_S(\mathbf{x})\text{)} \\
&\approx \int (\alpha^\intercal \mathbf{x})\, m(\mathbf{x}^\intercal \hat{\theta}) \mathcal{P}_S(\mathbf{x}) d\mathbf{x} && \text{(By assumption)} \\
&\approx \alpha^\intercal \frac{1}{N_S} \sum_{i=1}^{N_S} \mathbf{x}_i m(\mathbf{x}_i^\intercal \hat{\theta}) && \text{(Law of large numbers)}
\end{aligned}
$$

Approximately equivalent weights for (some) logistic regression MrP

- Suppose the model is $m(\mathbf{x}^\intercal \theta) = \text{Logistic}(\mathbf{x}^\intercal \theta)$, with MLE $\hat{\theta}$.
- MrP is $\hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T} \sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal \hat{\theta})$.

Example

Suppose $\frac{\mathcal{P}_T(\mathbf{x})}{\mathcal{P}_S(\mathbf{x})} \approx \alpha^\intercal \mathbf{x}$ for some α. **Then MrP is a *approximately* a CW estimator.**

$$
\begin{aligned}
\hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}}) &= \frac{1}{N_T} \sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal \hat{\theta}) \\
&\approx \int m(\mathbf{x}^\intercal \hat{\theta}) \mathcal{P}_T(\mathbf{x}) d\mathbf{x} && \text{(Law of large numbers)} \\
&= \int \frac{\mathcal{P}_T(\mathbf{x})}{\mathcal{P}_S(\mathbf{x})} m(\mathbf{x}^\intercal \hat{\theta}) \mathcal{P}_S(\mathbf{x}) d\mathbf{x} && \text{(Multiply by } \mathcal{P}_S(\mathbf{x})/\mathcal{P}_S(\mathbf{x})\text{)} \\
&\approx \int (\alpha^\intercal \mathbf{x})\, m(\mathbf{x}^\intercal \hat{\theta}) \mathcal{P}_S(\mathbf{x}) d\mathbf{x} && \text{(By assumption)} \\
&\approx \alpha^\intercal \frac{1}{N_S} \sum_{i=1}^{N_S} \mathbf{x}_i m(\mathbf{x}_i^\intercal \hat{\theta}) && \text{(Law of large numbers)} \\
&= \alpha^\intercal \frac{1}{N_S} \sum_{i=1}^{N_S} \mathbf{x}_i y_i && \text{(Property of exponential family MLEs)}
\end{aligned}
$$

Nearly equivalent weights for (some) logistic regression MrP

- Suppose the model is $m(\mathbf{x}^\intercal \theta) = \mathrm{Logistic}(\mathbf{x}^\intercal \theta)$, with MLE $\hat{\theta}$.
- MrP is $\hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T} \sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal \hat{\theta})$.

Example

Suppose $\frac{\mathcal{P}_T(\mathbf{x})}{\mathcal{P}_S(\mathbf{x})} \approx \alpha^\intercal \mathbf{x}$ for some α. **Then MrP is a *approximately* a CW estimator.**

$$\hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T} \sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal \hat{\theta}) = \frac{1}{N_S} \sum_{i=1}^{N_S} \underbrace{w_i^{\mathrm{MrP}}}_{\alpha^\intercal \mathbf{x}_i} y_i + \text{Small error}$$

[3] Krantz and Parks 2012; **G.**, Stephenson, et al. 2019.

Nearly equivalent weights for (some) logistic regression MrP

- Suppose the model is $m(\mathbf{x}^\intercal \theta) = \text{Logistic}(\mathbf{x}^\intercal \theta)$, with MLE $\hat{\theta}$.
- MrP is $\hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T} \sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal \hat{\theta})$.

Example

Suppose $\frac{\mathcal{P}_T(\mathbf{x})}{\mathcal{P}_S(\mathbf{x})} \approx \alpha^\intercal \mathbf{x}$ for some α. **Then MrP is a *approximately* a CW estimator.**

$$\hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T} \sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal \hat{\theta}) = \frac{1}{N_S} \sum_{i=1}^{N_S} \underbrace{w_i^{\text{MrP}}}_{\alpha^\intercal \mathbf{x}_i} y_i + \text{Small error}$$

But what are the weights? We don't observe $\frac{\mathcal{P}_T(\mathbf{x})}{\mathcal{P}_S(\mathbf{x})}$, so can't estimate α directly.

[3] Krantz and Parks 2012; **G.**, Stephenson, et al. 2019.

Nearly equivalent weights for (some) logistic regression MrP

- Suppose the model is $m(\mathbf{x}^\intercal \theta) = \text{Logistic}(\mathbf{x}^\intercal \theta)$, with MLE $\hat{\theta}$.
- MrP is $\hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T} \sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal \hat{\theta})$.

Example

Suppose $\frac{\mathcal{P}_T(\mathbf{x})}{\mathcal{P}_S(\mathbf{x})} \approx \alpha^\intercal \mathbf{x}$ for some α. **Then MrP is a *approximately* a CW estimator.**

$$\hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T} \sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal \hat{\theta}) = \frac{1}{N_S} \sum_{i=1}^{N_S} \underbrace{w_i^{\text{MrP}}}_{\alpha^\intercal \mathbf{x}_i} y_i + \text{Small error}$$

But what are the weights? We don't observe $\frac{\mathcal{P}_T(\mathbf{x})}{\mathcal{P}_S(\mathbf{x})}$, so can't estimate α directly.

Key idea (informal)

If $\hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}})$ is approximately linear, then $w_i^{\text{MrP}} \approx N_S \frac{\partial \hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}})}{\partial y_i}$.

[3] Krantz and Parks 2012; **G.**, Stephenson, et al. 2019.

Nearly equivalent weights for (some) logistic regression MrP

- Suppose the model is $m(\mathbf{x}^\intercal \theta) = \text{Logistic}(\mathbf{x}^\intercal \theta)$, with MLE $\hat{\theta}$.
- MrP is $\hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T} \sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal \hat{\theta})$.

Example

Suppose $\frac{\mathcal{P}_T(\mathbf{x})}{\mathcal{P}_S(\mathbf{x})} \approx \alpha^\intercal \mathbf{x}$ for some α. **Then MrP is a *approximately* a CW estimator.**

$$\hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T} \sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal \hat{\theta}) = \frac{1}{N_S} \sum_{i=1}^{N_S} \underbrace{w_i^{\text{MrP}}}_{\alpha^\intercal \mathbf{x}_i} y_i + \text{Small error}$$

But what are the weights? We don't observe $\frac{\mathcal{P}_T(\mathbf{x})}{\mathcal{P}_S(\mathbf{x})}$, so can't estimate α directly.

Key idea (informal)

If $\hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}})$ is approximately linear, then $w_i^{\text{MrP}} \approx N_S \frac{\partial \hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}})}{\partial y_i}$.

For logistic regression, compute and analyze $\frac{\partial \hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}})}{\partial y_i}$ using the implicit function theorem.[3]

[3] Krantz and Parks 2012; **G.**, Stephenson, et al. 2019.

Locally equivalent weights for hierarchical logistic regression MrP

- Suppose the model is $m(\mathbf{x}^\intercal \theta) = \mathrm{Logistic}(\mathbf{x}^\intercal \theta)$.
- Set a hierarchical prior $\mathcal{P}(\theta|\Sigma)\mathcal{P}(\Sigma)$, use MCMC to draw from $\mathcal{P}(\theta|\text{Survey data})$.
- MrP is $\hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T} \sum_{j=1}^{N_T} \mathbb{E}_{\mathcal{P}(\theta|\text{Survey data})} \left[m(\mathbf{x}_j^\intercal \theta) \right]$.

No reason to think $Y_{\mathcal{S}} \mapsto \hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}})$ is even approximately **globally** linear.

[4] Diaconis and Freedman 1986; Gustafson 1996; Efron 2015; **G.**, Broderick, and Jordan 2018.

Locally equivalent weights for hierarchical logistic regression MrP

- Suppose the model is $m(\mathbf{x}^\intercal\theta) = \mathrm{Logistic}(\mathbf{x}^\intercal\theta)$.
- Set a hierarchical prior $\mathcal{P}(\theta|\Sigma)\mathcal{P}(\Sigma)$, use MCMC to draw from $\mathcal{P}(\theta|\text{Survey data})$.
- MrP is $\hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T}\sum_{j=1}^{N_T}\mathbb{E}_{\mathcal{P}(\theta|\text{Survey data})}\left[m(\mathbf{x}_j^\intercal\theta)\right]$.

No reason to think $Y_{\mathcal{S}} \mapsto \hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}})$ is even approximately **globally** linear.

But we can still compute and analyze $w_i^{\mathrm{MrP}} := N_S \frac{\partial \hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}})}{\partial y_i}$ using Bayesian sensitivity analysis![4]

MrP weights for MCMC

$$w_i^{\mathrm{MrP}} := N_S \frac{\partial \hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}})}{\partial y_i} = N_S \frac{1}{N_T}\sum_{j=1}^{N_T}\underbrace{\mathrm{Cov}_{\mathcal{P}(\theta|\text{Survey data})}\left(m(\mathbf{x}_j^\intercal\theta), \theta^\intercal\mathbf{x}_i\right)}_{\text{Can estimate without rerunning MCMC!}}$$

[4] Diaconis and Freedman 1986; Gustafson 1996; Efron 2015; **G.**, Broderick, and Jordan 2018.

Locally equivalent weights for hierarchical logistic regression MrP

- Suppose the model is $m(\mathbf{x}^\intercal\theta) = \mathrm{Logistic}(\mathbf{x}^\intercal\theta)$.
- Set a hierarchical prior $\mathcal{P}(\theta|\Sigma)\mathcal{P}(\Sigma)$, use MCMC to draw from $\mathcal{P}(\theta|\text{Survey data})$.
- MrP is $\hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T}\sum_{j=1}^{N_T}\mathbb{E}_{\mathcal{P}(\theta|\text{Survey data})}\left[m(\mathbf{x}_j^\intercal\theta)\right]$.

No reason to think $Y_{\mathcal{S}} \mapsto \hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}})$ is even approximately **globally** linear.

But we can still compute and analyze $w_i^{\mathrm{MrP}} := N_{\mathcal{S}}\frac{\partial\hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}})}{\partial y_i}$ using Bayesian sensitivity analysis![4]

MrP weights for MCMC

$$w_i^{\mathrm{MrP}} := N_{\mathcal{S}}\frac{\partial\hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}})}{\partial y_i} = N_{\mathcal{S}}\frac{1}{N_T}\sum_{j=1}^{N_T}\underbrace{\mathrm{Cov}_{\mathcal{P}(\theta|\text{Survey data})}\left(m(\mathbf{x}_j^\intercal\theta), \theta^\intercal\mathbf{x}_i\right)}_{\text{Can estimate without rerunning MCMC!}}$$

What do these weights mean? There are now two distinct possibilities:

- "Locally implicit weights"
  - An estimator of $\mathcal{P}_T(\mathbf{x})/\mathcal{P}_{\mathcal{S}}(\mathbf{x})$ (via Riesz regression applied to the Gateaux derivative)
- "Locally equivalent weights"
  - A characterization of $Y_{\mathcal{S}} \mapsto \hat{\mu}^{\mathrm{MrP}}(Y_{\mathcal{S}})$ for diagnostics and interpretation

[4] Diaconis and Freedman 1986; Gustafson 1996; Efron 2015; **G.**, Broderick, and Jordan 2018.

Locally equivalent weights for hierarchical logistic regression MrP

- Suppose the model is $m(\mathbf{x}^\intercal\theta) = \text{Logistic}(\mathbf{x}^\intercal\theta)$.
- Set a hierarchical prior $\mathcal{P}(\theta|\Sigma)\mathcal{P}(\Sigma)$, use MCMC to draw from $\mathcal{P}(\theta|\text{Survey data})$.
- MrP is $\hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T}\sum_{j=1}^{N_T} \mathbb{E}_{\mathcal{P}(\theta|\text{Survey data})}\left[m(\mathbf{x}_j^\intercal\theta)\right]$.

No reason to think $Y_{\mathcal{S}} \mapsto \hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}})$ is even approximately **globally** linear.

But we can still compute and analyze $w_i^{\text{MrP}} := N_S \frac{\partial \hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}})}{\partial y_i}$ using Bayesian sensitivity analysis![4]

MrP weights for MCMC

$$w_i^{\text{MrP}} := N_S \frac{\partial \hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}})}{\partial y_i} = N_S \frac{1}{N_T}\sum_{j=1}^{N_T} \underbrace{\text{Cov}_{\mathcal{P}(\theta|\text{Survey data})}\left(m(\mathbf{x}_j^\intercal\theta), \theta^\intercal\mathbf{x}_i\right)}_{\text{Can estimate without rerunning MCMC!}}$$

What do these weights mean? There are now two distinct possibilities:

- "Locally implicit weights"
  - An estimator of $\mathcal{P}_T(\mathbf{x})/\mathcal{P}_S(\mathbf{x})$ (via Riesz regression applied to the Gateaux derivative)
- "Locally equivalent weights" $\leftarrow$ **The present talk will focus on this interpretation**
  - A characterization of $Y_{\mathcal{S}} \mapsto \hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}})$ for diagnostics and interpretation

[4] Diaconis and Freedman 1986; Gustafson 1996; Efron 2015; **G.**, Broderick, and Jordan 2018.

Locally equivalent weights for hierarchical logistic regression MrP

- Suppose the model is $m(\mathbf{x}^\intercal \theta) = \text{Logistic}(\mathbf{x}^\intercal \theta)$.
- Set a hierarchical prior $\mathcal{P}(\theta|\Sigma)\mathcal{P}(\Sigma)$, use MCMC to draw from $\mathcal{P}(\theta|\text{Survey data})$.
- MrP is $\hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}}) = \frac{1}{N_T} \sum_{j=1}^{N_T} \mathbb{E}_{\mathcal{P}(\theta|\text{Survey data})} \left[m(\mathbf{x}_j^\intercal \theta) \right]$.

MrP locally equivalent weights (MrPlew)

For new data $\tilde{Y}_{\mathcal{S}}$, form a **MrP locally equivalent weighting**:

$$\hat{\mu}^{\text{MrP}}(\tilde{Y}_{\mathcal{S}}) \approx \hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}}) + \sum_{i=1}^{N_S} w_i^{\text{MrP}} (\tilde{y}_i - y_i) \quad \text{where} \quad w_i^{\text{MrP}} := \frac{\partial \hat{\mu}^{\text{MrP}}(Y_{\mathcal{S}})}{\partial y_i}.$$

Our task is to rigorously show that even such local weights can be meaningfully used diagnostically in the same ways we use global weights.

Future work

Notice that there was no discussion of misspecification!

Calibration weights (typically) do not depend on $Y_{\mathcal{S}}$.

Future work

Notice that there was no discussion of misspecification!

Calibration weights (typically) do not depend on $Y_{\mathcal{S}}$.

But the high level idea can be extended much more widely:

1. Assume your initial model was accurate
2. Select some perturbation your model should be able to capture
3. Use local sensitivity to detect whether the change is what you expect
4. If the change is not what you expect, either (1) or (2) was wrong

Future work

Notice that there was no discussion of misspecification!

Calibration weights (typically) do not depend on $Y_{\mathcal{S}}$.

But the high level idea can be extended much more widely:

1. Assume your initial model was accurate
2. Select some perturbation your model should be able to capture
3. Use local sensitivity to detect whether the change is what you expect
4. If the change is not what you expect, either (1) or (2) was wrong

Such checks recover generlized versions of many standard diagnostics for linear models.

Examples:

- Additive parameter shifts $\leftrightarrow$ Unbiasedness
- Invariance to survey data weighting $\leftrightarrow$ Regressor + residual orthogonality
- Importance sampling $\leftrightarrow$ Sandwich covariance $\stackrel{?}{=}$ Inverse Fisher information

Future work

Student contributions and ongoing work:

- **Vladimir Palmin** is working on extending MrPlew to `lme4`
- **Sequoia Andrade** is working on generalizing to other local sensitivity checks
- **Lucas Schwengber** is working on novel flow–based techniques for local sensitivity
- **(Currently recruiting!)** Doubly–robust Bayesian Hierarchical MrP

Vladimir Palmin

Sequoia Andrade

Lucas Schwengber

Preprint and R package (hopefully) coming soon!

References i

B., Eli, Avi F., and Erin H. (2021). *Multilevel calibration weighting for survey data*. arXiv: 2102.09052 [stat.ME].

Chattopadhyay, A. and J. Zubizarreta (2023). "On the implied weights of linear regression for causal inference". In: *Biometrika* 110.3, pp. 615–629.

Diaconis, P. and D. Freedman (1986). "On the consistency of Bayes estimates". In: *The Annals of Statistics*, pp. 1–26.

Efron, B. (2015). "Frequentist accuracy of Bayesian estimates". In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 77.3, pp. 617–646.

G., T. Broderick, and M. I. Jordan (2018). "Covariances, robustness and variational bayes". In: *Journal of machine learning research* 19.51.

G., W. Stephenson, et al. (2019). "A swiss army infinitesimal jackknife". In: *The 22nd International Conference on Artificial Intelligence and Statistics*. PMLR, pp. 1139–1147.

Gelman, A. (2007a). "Rejoinder: Struggles with survey weighting and regression modelling". In: *Statistical Science* 22.2, pp. 184–188.

— (2007b). "Struggles with survey weighting and regression modeling". In.

Gustafson, P. (1996). "Local sensitivity of posterior expectations". In: *The Annals of Statistics* 24.1, pp. 174–195.

Krantz, S. and H. Parks (2012). *The Implicit Function Theorem: History, Theory, and Applications*. Springer Science & Business Media.